

ABSTRACT:

Hashes are short summaries or signatures of data files which can be used to identify the file. Hashing multimedia content (audio, video, images) is difficult because the hash of original content and processed (e.g. compressed) content may differ significantly.

The disclosed method generates robust hashes for multimedia content, for example, audio clips. The audio clip is divided (12) into successive (preferably overlapping) frames. For each frame, the frequency spectrum is divided (15) into bands. A robust property of each band (e.g. energy) is computed (16) and represented (17) by a respective hash bit. An audio clip is thus represented by a concatenation of binary hash words, one for each frame. To identify a possibly compressed audio signal, a block of hash words derived therefrom is matched by a computer (20) with a large database (21). Such matching strategies are also disclosed. In an advantageous embodiment, the extraction process also provides information (19) as to which of the hash bits are the least reliable. Flipping these bits considerably improves the speed and performance of the matching process.

Fig. 1.